What is claimed is:

100
4
14 10
LT
The Tien
4.4
I
122
iž
rek
IJ
I.J
١,٠٠٠
125

3

4

5

6

7

8

9

1

2

1

2

3

1\	A computing system,	comprising
----	---------------------	------------

a rounding apparatus to accepts an input value that is a real number represented in floating-point format, and to perform a rounding operation on the input value to generate an output value that is an integer represented in floating-point format;

CLAIMS

a memory to store a computer program that utilizes the rounding apparatus; and

a central processing unit (CPU) to execute the computer program, the CPU is cooperatively connected to the rounding apparatus and the memory.

- 2. The system of claim 1, wherein the rounding apparatus uses a truncation technique to round the input value.
- 3. The system of claim 2, wherein the rounding apparatus includes: a floating-point to integer converter to truncate the input value to convert the input value to an integer represented in an integer format; and
- an integer to floating-point converter to convert the integer represented in an integer format to the output value.
- 1 4. The system of claim 1, wherein the rounding apparatus rounds the 2 input value to the nearest integer.
- The system of claim 4, wherein the rounding apparatus includes:
- an "AND" operator to extract a sign bit of the input value;
- an "OR" operator to generate an adjustment value based on the sign bit;

-22-

42390P10415

PAT. APPL.

2

3

1

2

3

1

2

1

2

3

N	Λ
4)	an ADD operator to compute an adjusted input value by adding the
5 (adjustment value to the input value, the adjusted input value is a real number
6	represented in floating-point format;
7	a floating-point to integer converter to truncate a fractional portion of the
8	adjusted input value to convert the adjusted input value to an integer
9	represented in an integer format; and
10	an integer to floating-point converter to convert the integer represented in
11	an integer format to generate the output value.

- 6. The system of claim 5, wherein the "AND" operator extracts the sign bit of the input value by performing a bit-wise logical AND operation on the input value and a sign mask.
- 7. The system of claim 5, wherein the "OR" operator generates the adjustment value by performing a bit-wise logical OR operation on the sign bit and a real value of 0.5.
- 8. The system of claim 1, wherein the rounding apparatus rounds the input value toward minus infinity $(-\infty)$.
- 9. The system of claim 8, wherein the rounding apparatus includes: a floating-point to integer converter to truncate an input value to convert the input value to a first integer represented in an integer format;
- an integer to floating-point converter to convert the first integer
 represented in an integer format to a second integer represented in floating-point
 format;
- a first SUBTRACT operator to compute a fractional portion of the input value using the second integer;

42390P10415 -23- PAT. APPL.

	VI				
D		n "less than" comparator to generate a boolean mask based on the			
7	10(fractional portion of the input value;			
t	11	an ("AND" operator to use the boolean mask to generate an adjustment			
	12	value represented in floating-point format; and			
	13	a second SUBTRACT operator to subtract the adjustment value from the			
	14	input value to generate the output value.			
	1	10. The system of claim 9, wherein the first SUBTRACT operator			
	2	computes the fractional portion of the input value by subtracting the second			
1 ····	3	integer from the input value.			
	1	11. The system of claim 9, wherein the "less than" comparator			
	2	generates the boolean mask by comparing the fractional portion of the input			
Land Street Street Street Street Street Street Street	3	value to a real value of 0.0.			
H					
1	1	12. The system of claim wherein the "AND" operator generates the			
	2	adjustment value by performing a bit wise logical AND operation on the boolear			
that the time that	3	mask and a real value of 1.0.			
بَعدا					
	1	13. The system of claim 1, wherein the rounding apparatus rounds the			
	2	input value toward plus infinity $(+\infty)$.			
	1	14. The system of claim 13, wherein the rounding apparatus includes:			
	2	a floating-point to integer converter to truncate an input value to convert			
	3	the input value to a first integer represented in an integer format;			
	4	an integer to floating-point converter to convert the first integer			

42390P10415 -24- PAT. APPL.

represented in an integer format to a second integer represented in floating-point

5

6

format;

#	0.70	a SUBTRACT operator to compute a fractional portion of the input value		
ı	8	using the second integer;		
9 a "greater-than" comparator to generate a boolean mask based				
	10	fractional portion of the input value;		
an "ANO" operator to use the boolean mask to generate an adju				
	12	value represented in floating-point format; and		
	13	an ADD operator to add the adjustment value to the input value to		
14 generate the output value.				
	1	15. The system of claim 14, wherein the SUBTRACT operator computes		
	2	the fractional portion of the input value by subtracting the second integer from		
. Three street Burne. Th	3	the input value.		
Med and thus Au	1	16. The system of claim 4, wherein the "greater-than" comparator		
Ť	2	generates the boolean mask by comparing the fractional portion of the input		
Spart Lend their time there I'm	3	value to a real value of 0.0.		
# T	1	17. The system of claim 14, wherein the "AND" operator generates the		
1	2	adjustment value by performing a bit-wise logical AND operation on the boolean		
	3	mask and a real value of 1.0.		
	1	18. A method comprising:		
	2	accepting an input value that is a real number represented in floating-		
	3	point format;		
	4	converting the input value to a first integer;		
	5	converting the first integer represented to a second integer; and		
	6	storing the second integer as an output value.		

42390P10415 -25- PAT. APPL.

. .

	1	19. The method of claim 18, wherein converting the input value to a
,	P	first integer comprises:
	3	representing the first integer in an integer format.
	1	20. The method of claim 18, wherein converting the first integer to the
	2	second integer comprises:
	3	representing the second integer in floating-point format.
	1	21. A method comprising:
	2	building an adjustment value represented in floating-point format;
	3	adding the adjustment value to an input value to generate an adjusted
	4	input value represented in floating-point format;
	5	truncating the adjusted input value to convert the adjusted input value to
	6	a first integer represented in an integer format;
	7	converting the first integer represented in an integer format to a second
	8	integer represented in floating-point format; and
	9	storing the second integer as an output value.
	1	22. The method of claim 21, wherein building the adjustment value
	2	comprises:
	3	extracting a sign bit of the input value by performing a bit-wise logical
	4	AND operation on the input value and a sign mask.
	1	23. The method of claim 21, wherein building the adjustment value
	2	comprises:
	3	building the adjustment value by performing a bit-wise logical OR
	4	operation on a real value of 0.5 and a sign bit extracted from the input value.
		\

42390P10415 -26- PAT. APPL.

ſ	
1/2	24. A method comprising:
\$	24. A method comprising: generating a first integer represented in an integer format by truncating
3	an input value;
4	converting the first integer represented in an integer format to a second
5	integer represented in floating-point format;
6	computing a fractional portion of the input value using the second integer
7	represented in floating-point format;
8	generating aboolean value using the fractional portion of the input value;
9	creating an adjustment value using the boolean value;
10	computing a rounded input value by subtracting the adjustment value
11	from the input value.
1	25. The method of claim 24, wherein computing the fractional portion
2	of the input value comprises:
3	subtracting the second integer represented in floating-point format from
4	the input value to generate the ractional portion of the input value.
1	26. The method of claim 24, wherein generating the boolean value
2	comprises comparing the fractional portion of the input value to a real value of
3	0.0.
1	27. The method of claim 24, wherein creating an adjustment value
2	comprises performing a bit-wise logical AND operation on the boolean value
3	and a real value of 1.0.
1	28. A method comprising:
2	generating a first integer represented in an integer format by truncating
3	an input value;

42390P10415 -27- PAT. APPL.

1		Sonve	erting the first integer represented in an integer format to a second
		1	
5 0	intege	1	esented in floating-point format;
6		subtr	acting the second integer represented in floating-point format from
7	the in	put va	ue to generate a fractional portion of the input value;
8		gener	rating a boolean value using the fractional portion of the input value;
9		creati	ing an adjustment value using the boolean value;
10		addir	ng the adjustment value to the input value to generate a rounded
11	input	value.	
1		29.	The method of claim 28, wherein creating an adjustment value
2	comp	rises:	
3		comp	earing the fractional portion of the input value to a real value of 0.0.
1		30.	The method of claim 28, wherein creating an adjustment value
2	comp	rises:	
3		perfo	rming a bit-wise logical AND operation on the boolean value and a
4	real v	alue of	1.0.
1		31.	A machine-readable medium comprising instructions which, when
2	execu	ted by	a machine, cause the machine to perform operations comprising:
3		a first	code segment truncates the input value to convert the input value to
4	a first	intege	r; and
5		a seco	ond code segment integer to convert the first integer to a second
6	intege	er.	
1		32.	The machine-readable medium of claim 31, wherein the first integer

42390P10415 -28- PAT. APPL.

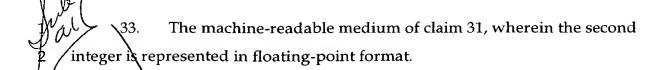
is represented in an integer format.

2

8

9

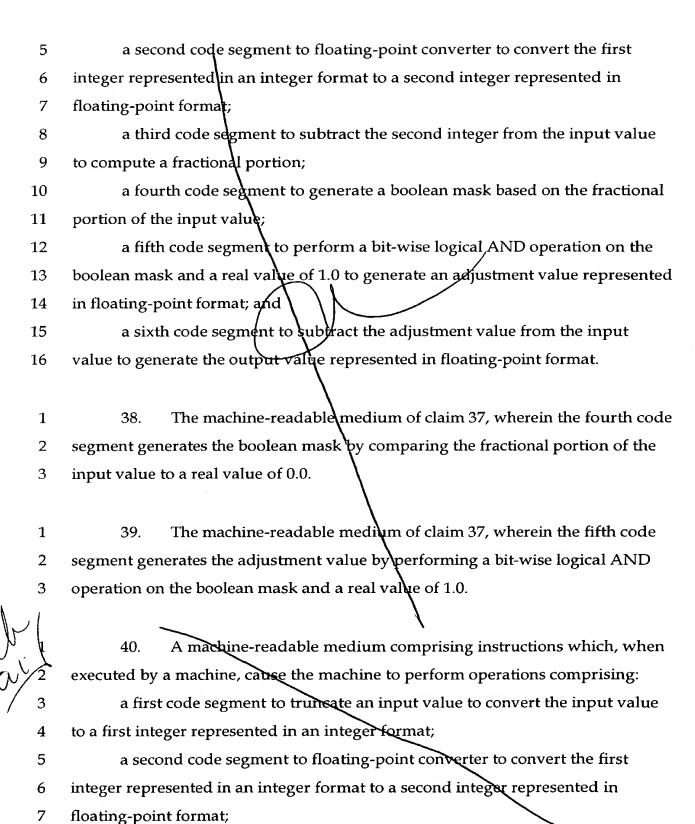
10



- 34. A machine-readable medium comprising instructions which, when executed by a machine, cause the machine to perform operations comprising:
- a first code segment to extract a sign bit of the input value;
- a second code segment to generate an adjustment value based on the sign 5 bit;
- a third code segment to compute an adjusted input value represented in floating-point format;
 - a fourth code segment to truncate a fractional portion of the adjusted input value to convert the adjusted input value to an integer represented in an integer format; and
- 11 a fifth code segment to convert the integer represented in an integer 12 format to generate the output value.
- 1 35. The machine-readable medium of claim 34, wherein the second 2 code segment generates the adjustment value by performing a bit-wise logical 3 OR operation on the sign bit and a value of 0.5.
- 1 36. The machine-readable medium of claim 34, wherein the third code 2 segment computes the adjusted input value by adding the adjustment value to 3 the input value.
- 37. A machine-readable medium comprising instructions which, when executed by a machine, cause the machine to perform operations comprising:

 a first code segment to truncate an input value to convert the input value
- 4 to a first integer represented in an integer format;

42390P10415 -29- PAT. APPL.



-30-

42390P10415

PAT. APPL.



12

13

14

15

1

2

3

1

2

3

a third code segment to subtract the second integer from the input value to compute a fractional portion of the input value;

a fourth code segment to generate a boolean mask based on the fractional portion of the input value;

a fifth code segment to an adjustment value represented in floating-point format; and

a sixth code segment to subtract the adjustment value from the input value to generate the output value represented in floating-point format.

- 41. The machine-readable medium of claim 40, wherein the fourth code segment generates the boolean mask by comparing the fractional portion of the input value to a real value of 0.0.
- 42. The machine-readable medium of claim 40, wherein the fifth code segment generates the adjustment value by performing a bit-wise logical AND operation on the boolean mask and a real value of 1.0.

